

Bell correlations in a many-body system with finite statistics

Sebastian Wagner,¹ Roman Schmied,² Matteo Fadel,² Philipp Treutlein,² Nicolas Sangouard,¹ and Jean-Daniel Bancal¹

¹*Quantum Optics Theory Group, Department of Physics, University of Basel, Klingelbergstrasse 82, 4056 Basel, Switzerland*

²*Quantum Atom Optics Lab, Department of Physics, University of Basel, Klingelbergstrasse 82, 4056 Basel, Switzerland*

(Dated: February 13, 2017)

A recent experiment reported the first violation of a Bell correlation witness in a many-body system [Science 352, 441 (2016)]. Following discussions in this paper, we address here the question of the statistics required to witness Bell correlated states, i.e. states violating a Bell inequality, in such experiments. We start by deriving multipartite Bell inequalities involving an arbitrary number of measurement settings, two outcomes per party and one- and two-body correlators only. Based on these inequalities, we then build up improved witnesses able to detect Bell-correlated states in many-body systems using two collective measurements only. These witnesses can potentially detect Bell correlations in states with an arbitrarily low amount of spin squeezing. We then establish an upper bound on the statistics needed to convincingly conclude that a measured state is Bell-correlated.

a. Introduction – Physics research fundamentally relies on the proper analysis of finite experimental data. In this exercise, assumptions play a subtle but crucial role. On the one hand, they are needed in order to reach a conclusion; even device-independent assessments rely on assumptions [1]. On the other hand, they open the door for undesirable effects ranging from a reduction of the conclusion's scope, when more assumptions are used than strictly needed, to biased results when relying on unmet assumptions. Contrary to popular belief, such cases are frequent in science, even for common assumptions [2, 3]. Relying on fewer hypotheses, when possible, is thus desirable to obtain more general, accurate and trustworthy conclusions [4, 5].

Bell nonlocality, as revealed by the violation of a Bell inequality, constitutes one of the strongest forms of non-classicality known today. However, its demonstration has long been restricted to systems involving few particles [6–10]. Recently, the discovery of multipartite Bell inequalities that only rely on one- and two-body correlators opened up new possibilities [11]. Although these inequalities have not yet lead to the realization of a multipartite Bell test, they can be used to derive witnesses able to detect Bell correlated states, i.e. states capable of violating a Bell inequality.

Using such a witness, an experiment recently detected the presence of Bell correlations in a many-body system under the assumption of gaussian statistics [12]. While this demonstration uses spin squeezed states, the detection of Bell correlations in other systems was also recently investigated [13]. The witness used in Ref. [12] involves one- and two-body correlation functions and takes the form $\mathcal{W} \geq 0$, where the inequality is satisfied by measurements on states that are not Bell-correlated. Observation of a negative value for \mathcal{W} then leads to the conclusion that the measured system is Bell-correlated. However, due to the statistics loophole [14, 15], reaching such a conclusion in the presence of finite statistics requires special care. In particular, an assessment of the

probability with which a non-Bell-correlated state could be responsible for the observed data is required before concluding about the presence of Bell correlations without further assumptions.

Concretely, the witness of Refs. [12] has the property of admitting a quantum violation lower-bounded by a constant $\mathcal{W}_{\text{opt}} < 0$, while the largest possible value $\mathcal{W}_{\text{max}} > 0$ is achievable by a product state and increases linearly with the size of the system N . These properties imply that a small number of measurements on a state of the form

$$\rho = (1 - q)|\psi\rangle\langle\psi| + q(|\uparrow\rangle\langle\uparrow|)^{\otimes N}, \quad (1)$$

where $\mathcal{W}(|\psi\rangle) = \mathcal{W}_{\text{opt}}$, $\mathcal{W}(|\uparrow\rangle^{\otimes N}) = \mathcal{W}_{\text{max}}$ and q is small, is likely to produce a negative estimate of \mathcal{W} , even though the state is not detected by the witness in the limit of infinitely many measurement rounds [12]. This state thus imposes a lower bound on the number of measurements required to exclude, through such witnesses, all non-Bell-correlated states with high confidence. Contrary to other assessments, this lower bound increases with the number of particles involved in the many-body system. Therefore, it is not captured by the standard deviation of one- and two-body correlation functions (which on the contrary decreases as the number of particles increases).

It is worth noting that states of the form (1) put similar bounds on the number of measurements required to perform any hypothesis tests in a many-body system satisfying the conditions above. This includes in particular tests of entanglement [16–19] based on the entanglement witnesses of Ref. [20–22].

In this article, we address this statistical problem in the case of Bell correlation detection by providing a number of measurement rounds sufficient to exclude non-Bell-correlated states from an observed witness violation. Let us mention that in Ref. [12], the statistics loophole is circumvented by the addition of an assumption on the set of local states being tested. This has the effect of

reducing the scope of the conclusion: the data reported in Ref. [12], are only able to exclude a subset of all non-Bell-correlated states (as pointed out in the reference). Here, we show that such additional assumptions are not required in experiments on many-body systems, and thus argue that they should be avoided in the future.

In order to minimize the amount of statistics required to reach our conclusion, we start by investigating improved Bell correlation witnesses. For this, we first derive Bell inequalities with two-body correlators and an arbitrary number of settings. This allows us to obtain Bell-correlation witnesses that are more resistant to noise compared to the one known to date [12]. We then analyse the statistical properties of these witnesses and provide an upper bound on the number of measurement rounds needed to rule out all local states in a many-body system. We show that this upper bound is linear in the number of particles, hence making the detection of Bell correlations free of the statistical loophole possible in systems with a large number of particles.

b. Symmetric two-body correlator Bell inequalities with an arbitrary number of settings – Multipartite Bell inequalities that are symmetric under exchange of parties and which involve only one- and two-body correlators have been proposed in scenarios where each party uses two measurement settings and receives an outcome among two possible results [11]. Similar inequalities were also obtained for translationally invariant systems [23], or based on Hamiltonians [24]. Here, we derive a similar family of Bell inequalities that is invariant under arbitrary permutations of parties but allows for an arbitrary number of measurement settings per party.

Let us consider a scenario in which N parties can each perform one of m possible measurements $M_k^{(i)}$ ($k = 0, \dots, m-1$; $i = 1, \dots, N$) with binary outcomes ± 1 . We write the following inequality:

$$I_{N,m} = \sum_{k=0}^{m-1} \alpha_k S_k + \frac{1}{2} \sum_{k,l} S_{kl} \geq -\beta_c, \quad (2)$$

where $\alpha_k = m - 2k - 1$, β_c is the local bound, and the symmetrized correlators are defined as

$$S_k := \sum_{i=1}^N \langle M_k^{(i)} \rangle, \quad S_{kl} := \sum_{i \neq j} \langle M_k^{(i)} M_l^{(j)} \rangle. \quad (3)$$

Let us show that (2) is a valid Bell inequality for $\beta_c = \left\lfloor \frac{m^2 N}{2} \right\rfloor$, where $\lfloor x \rfloor$ is the largest integer smaller or equal to x . Below, we assume that m is even; see Appendix A for the case of odd m .

Since $I_{N,m}$ is linear in the probabilities and local behaviors can be decomposed as a convex combination of deterministic local strategies, the local bound of Eq. (2) can be reached by a deterministic local strategy [25]. We thus restrict our attention to these strategies and write

$$\langle M_k^{(i)} \rangle = x_k^i = \pm 1 \quad \Rightarrow \quad S_{kl} = S_k S_l - \sum_{i=1}^N x_k^i x_l^i, \quad (4)$$

where x_k^i is the (deterministic) outcome party i produces when asked question k . This directly leads to the following decomposition:

$$I_{N,m} = \sum_{k=0}^{\frac{m}{2}-1} \alpha_k (S_k - S_{m-k-1}) + \frac{1}{2} B^2 - \frac{1}{2} C \geq -\beta_c, \quad (5)$$

with $B := \sum_{k=0}^{m-1} S_k$ and $C := \sum_{i=1}^N \left(\sum_{k=0}^{m-1} x_k^i \right)^2$. Due to the symmetry under exchange of parties of this Bell expression, it is convenient to introduce, following [11], variables counting the number of parties that use a specific deterministic strategy:

$$\begin{aligned} a_{j_1 < \dots < j_n} &:= \# \{ i \in \{1, \dots, N\} | x_k^i = -1 \text{ iff } k \in \{j_1, \dots, j_n\} \} \\ \bar{a}_{j_1 < \dots < j_n} &:= \# \{ i \in \{1, \dots, N\} | x_k^i = +1 \text{ iff } k \in \{j_1, \dots, j_n\} \} \\ n &\leq \frac{m}{2}, \quad \bar{a}_{j_1, \dots, j_{\frac{m}{2}}} \equiv 0, \end{aligned} \quad (6)$$

where $\#$ denotes the set cardinality. Since each party has to choose a strategy, the variables sum up to N :

$$\sum_{\text{all variables}} = \sum_{n=0}^{\frac{m}{2}} \sum_{j_1 < \dots < j_n} (a_{j_1 \dots j_n} + \bar{a}_{j_1 \dots j_n}) = N. \quad (7)$$

The correlators can now be expressed as

$$S_k = \sum_{n=0}^{\frac{m}{2}} \sum_{j_1 < \dots < j_n} (a_{j_1 \dots j_n} - \bar{a}_{j_1 \dots j_n}) y_k^{j_1 \dots j_n}, \quad (8)$$

with $y_k^{j_1 \dots j_n} = -1$ if $k \in \{j_1, \dots, j_n\}$, and $+1$ otherwise.

The first term of (5) concerns the difference between two correlators. Let us see how this term decomposes as a function of the number of indices present in its variables. From Eq. (8), it is clear that a variable with n indices only appears in the difference $S_k - S_l$ if $y_k^{j_1 \dots j_n} \neq y_l^{j_1 \dots j_n}$. But the corresponding strategy only has n differing outcomes and each correlator in this term only appears once, so a variable with n indices appears in at most n of these differences. Moreover, if it appears, it does so with a factor ± 2 . The coefficient in front of a variable with n indices in the first sum of (5) thus cannot be smaller than $-2 \sum_{k=0}^{n-1} \alpha_k = 2n(n-m)$.

The second term of (5) can be bounded as $B^2 \geq 0$, while the third one can be expressed as

$$C = \sum_{n=0}^{\frac{m}{2}} \sum_{j_1 < \dots < j_n} (a_{j_1 \dots j_n} + \bar{a}_{j_1 \dots j_n}) (m-2n)^2. \quad (9)$$

Putting everything together and using property (7), we arrive at

$$\begin{aligned} I_{N,m} &\geq \sum_{k=0}^{\frac{m}{2}-1} \alpha_k (S_k - S_{m-k-1}) - \frac{1}{2} C \\ &\geq -\frac{m^2}{2} \sum_{\text{all variables}} = -\frac{m^2 N}{2} = -\beta_c, \end{aligned} \quad (10)$$

which concludes the proof.

Note that this bound is achieved for $a_{01\dots\frac{m}{2}-1} = N$, i.e. when for each party exactly the first half of the measurements yields result -1 . Note also that the Bell inequality (2) does not reduce to Ineq. (6) of Ref. [11] when $m = 2$. Indeed, while none of these inequalities is a facet of the local polytope, the latter one is a facet of the symmetrized 2-body correlator local polytope [11, 26].

c. From Bell inequalities to Bell-correlation witnesses

– Let us now derive a set of Bell-correlation witnesses assuming a certain form for the measurement operators. Here, no assumptions are made on the measured state.

Following Ref. [12], we start from inequality (2) and introduce spin measurements along the axes \vec{d}_k , $k = 0, \dots, m-1$, as well as the collective spin observables \hat{S}_k :

$$M_k^{(i)} = \vec{d}_k \cdot \vec{\sigma}^{(i)}, \quad \hat{S}_k = \frac{1}{2} \sum_{i=1}^N M_k^{(i)}, \quad (11)$$

where $\vec{\sigma}$ is the Pauli vector acting on a spin- $\frac{1}{2}$ system. The correlators can be expressed in terms of these total spin observables and the measurement directions [11]:

$$S_k = 2\langle \hat{S}_k \rangle$$

$$S_{kl} = 2 \left[\langle \hat{S}_k \hat{S}_l \rangle + \langle \hat{S}_l \hat{S}_k \rangle \right] - N \vec{d}_k \cdot \vec{d}_l. \quad (12)$$

This defines the Bell operators

$$\hat{W}_{N,m} := 2 \sum_{k=0}^{m-1} \alpha_k \hat{S}_k + 2 \sum_{k,l} \hat{S}_k \hat{S}_l - \frac{N}{2} \sum_{k,l} \vec{d}_k \cdot \vec{d}_l + \left[\frac{m^2 N}{2} \right], \quad (13)$$

whose expectation values are positive for states that are not Bell-correlated. Note that the expectation value of these operators need not be negative for all Bell-correlated states and every choice of measurement, though. A negative value may only be achieved for specific choices of states and measurement settings.

We now consider measurement directions $\vec{d}_k = \vec{a} \cos(\vartheta_k) + \vec{b} \sin(\vartheta_k)$ lying in a plane spanned by two orthonormal vectors \vec{a} and \vec{b} , with the antisymmetric angle distribution $\vartheta_{m-k-1} = -\vartheta_k$. Note that the coefficients α_k share the same antisymmetry. Defining $\mathcal{W}_m := \langle \frac{\hat{W}_{N,m}}{2N} \rangle$ for even m , we arrive at the following family of witnesses:

$$\mathcal{W}_m = C_b \sum_{k=0}^{\frac{m}{2}-1} \alpha_k \sin(\vartheta_k) - (1 - \zeta_a^2) \left[\sum_{k=0}^{\frac{m}{2}-1} \cos(\vartheta_k) \right]^2 + \frac{m^2}{4}, \quad (14)$$

with $\mathcal{W}_m \geq 0$ for states that are not Bell correlated. These Bell correlation witnesses depend on $\frac{m}{2}$ angles ϑ_k and involve just two quantities to be measured: the scaled collective spin $C_b := \langle \frac{\hat{S}_b}{N/2} \rangle$ and the scaled second moment $\zeta_a^2 := \langle \frac{\hat{S}_a^2}{N/4} \rangle$.

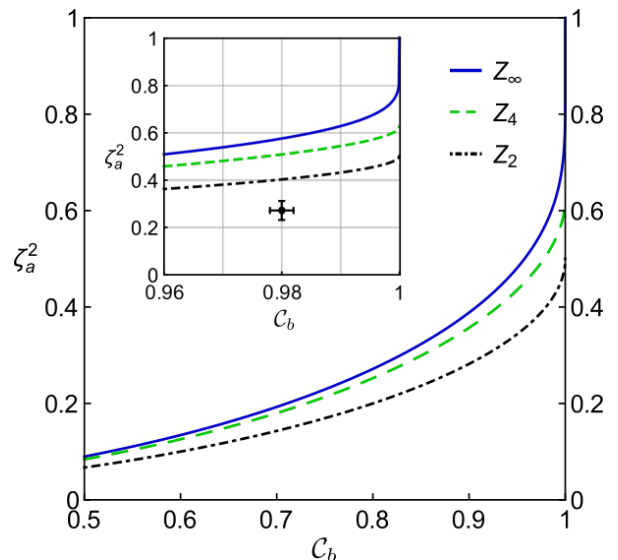


FIG. 1: Plots of the critical lines Z_2 , Z_4 and Z_∞ . The witness obtained from the Bell inequality with 4 settings already provides a significant improvement over the case of 2 settings. The black point in the inset shows the data point from [12], with $N = 476 \pm 21$.

The tightest constraints on C_b and ζ_a^2 that allow for a violation of $\mathcal{W}_m \geq 0$ are obtained by minimizing \mathcal{W}_m over the angles ϑ_k . Solving $\frac{\partial \mathcal{W}_m}{\partial \vartheta_k} = 0$ yields (see Appendix B):

$$\vartheta_k = -\arctan[\lambda_m(m-2k-1)], \quad (15)$$

$$\frac{C_b}{2\lambda_m(1-\zeta_a^2)} = \sum_{k=0}^{\frac{m}{2}-1} \cos(\vartheta_k). \quad (16)$$

Equation (16) is a self-consistency equation for λ_m that has to be satisfied in order to minimize \mathcal{W}_m .

Using these parameters, we can rewrite our witness in terms of the physical parameters C_b and ζ_a^2 only. For two measurement directions ($m = 2$), we find that states which are not Bell-correlated satisfy

$$\zeta_a^2 \geq Z_2(C_b) = \frac{1}{2} \left(1 - \sqrt{1 - C_b^2} \right). \quad (17)$$

This recovers the bound obtained from a different inequality in [12]. Note that in the present case, the argument is more direct since it does not involve C_a , the first moment of the spin operator in the a direction.

Increasing the number of measurement directions allows for the detection of Bell correlations in additional states. In the limit $m \rightarrow \infty$, we find (see Appendix B):

$$\zeta_a^2 \geq Z_\infty(C_b) = 1 - \frac{C_b}{\text{artanh}(C_b)}. \quad (18)$$

Figure 1 shows the two witnesses (17) and (18) together with the one obtained similarly for $m = 4$ settings in the C_b - ζ_a^2 plane. The curve Z_∞ reaches the point

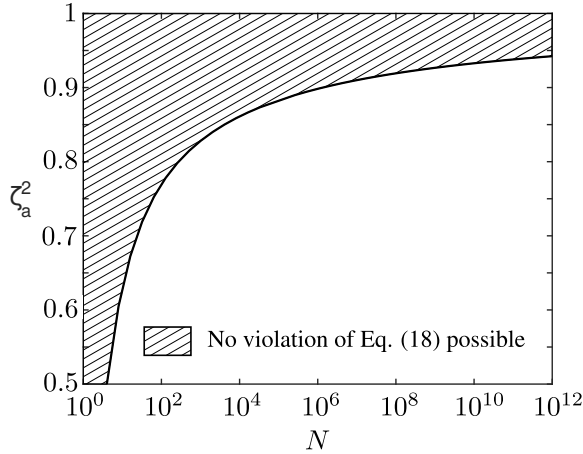


FIG. 2: Upper bound on the value of ζ_a^2 required to see a violation of the Bell correlation witness (18). The bound depends on the number of particles N .

$C_b = \zeta_a^2 = 1$, therefore allowing in principle for the detection of Bell correlations in presence of arbitrarily low squeezing. It is known, however, that some values of C_b and ζ_a^2 can only be reached in the limit of a large number of spins [27]. For any fixed N , a finite amount of squeezing is thus necessary in order to allow for the violation of our witness (see Appendix C). The corresponding upper bound on ζ_a^2 is shown in Figure 2.

Points below the curve Z_m in Fig. 1 indicate a violation of the witness $\mathcal{W}_m \geq 0$ obtained from the corresponding m -settings Bell inequality. Violation of any such bound reveals the presence of a Bell-correlated state. However, as discussed in the introduction, conclusions in the presence of finite statistics have to be examined carefully, since in practice, one can never conclude from the violation of a witness that the measured state is Bell correlated with 100% confidence. The point shown in the inset of Fig. 1 corresponds to the data reported in Ref. [12] from measurements on a spin-squeezed Bose-Einstein condensate. This point clearly violates the witnesses for $m = 2, 4, \infty$ by several standard deviations, although the number of measurement rounds is too small to guarantee that the measured state is Bell correlated without further assumptions [12].

d. Finite Statistics – In this section, we put a bound on the number of experimental runs needed to exclude with a given confidence that a measured state is not Bell-correlated. Note that such a conclusion does not follow straightforwardly from the violation of the witness by a fixed number of standard deviations. Indeed, standard deviations inform on the precision of a violation, but fail at excluding arbitrary local models [15], including e.g. models which may show non-gaussian statistics with rare events. We thus look here for a number of experimental runs that is sufficient to guarantee a p-value lower than a given threshold for the null hypothesis ‘The measured state is not Bell-correlated’. Since we are concerned

with the characterization of physical systems in the absence of an adversary, we assume that the same state is prepared in each round (i.i.d. assumption).

For this statistical analysis, let us consider a different Bell correlation witness than (18). Indeed, we derived this inequality in order to maximize the amount of violation for given data, but here we rather wish to maximize the statistical evidence of a violation. For this, we take (14) and consider the representation of the angles given in Eq. (15), but without taking Eq. (16) into account. In the limit of infinitely-many measurement settings, we find (see Appendix B)

$$\mathcal{W}_{\text{stat}} = -C_b \Delta_\nu - (1 - \zeta_a^2) \Lambda_\nu^2 + \frac{1}{4} \geq 0, \text{ with} \quad (19)$$

$$\Delta_\nu = \frac{\sqrt{1 + \nu^2}}{4\nu} - \frac{\text{arsinh}(\nu)}{4\nu^2}, \quad \Lambda_\nu = \frac{\text{arsinh}(\nu)}{2\nu}, \quad (20)$$

where $\nu = \lim_{m \rightarrow \infty} \lambda_m \cdot m$ is a free parameter that fully specifies the set of measurement angles.

In order to model the experimental evaluation of $\mathcal{W}_{\text{stat}}$, we introduce the following estimator:

$$\mathcal{T} = \frac{\chi(Z=0)}{q} X + \frac{\chi(Z=1)}{1-q} Y + \left(\frac{1}{4} - \Delta_\nu - \Lambda_\nu^2\right). \quad (21)$$

Here, χ denotes the indicator function and the binary random variable Z accounts for the choice between the measurement of either C_b or ζ_a . Each measurement allows for the evaluation of the corresponding random variables $X = \Delta_\nu(1 - C_b)$ and $Y = \Lambda_\nu^2 \zeta_a^2$. Assuming that Z is independent of X and Y and choosing $q = P[Z=0]$ guarantees that \mathcal{T} is a proper estimator of $\mathcal{W}_{\text{stat}}$, i.e. $\langle \mathcal{T} \rangle = \mathcal{W}$. q then corresponds to the probability of performing a measurement along the b axis. We choose $q = \left(1 + \frac{\Lambda_\nu^2 N}{2\Delta_\nu}\right)^{-1}$ so that the contributions of both measurements to \mathcal{T} have the same magnitude, i.e. the maximum values of X/q and $Y/(1-q)$ are equal within the domain $|C_b| \leq 1$ and $\zeta_a^2 \in [0, N]$. This also guarantees that the spectrum of \mathcal{T} matches that of $\mathcal{W}_{\text{stat}}$.

Suppose the measured state is non-Bell-correlated, i.e. that its mean value $\mu = \langle \mathcal{T} \rangle = \mathcal{W}_{\text{stat}} \geq 0$. We are now interested in the probability that after M experimental runs the estimated value $T = \frac{1}{M} \sum_{i=1}^M \mathcal{T}_i$ of the witness $\mathcal{W}_{\text{stat}}$ falls below a certain value $t_0 < 0$, with \mathcal{T}_i being the value of the estimator in the i^{th} run.

In statistics, concentration inequalities deal with exactly this issue. In Appendix D, we compare four of these inequalities, namely the Chernoff, Bernstein, Uspensky and Berry-Esseen ones and show explicitly that in the regime of interest the tightest and therefore preferred bound results from the Bernstein inequality:

$$P[T \leq t_0] \leq \exp \left(-\frac{(\mu - t_0)^2 M}{2\sigma_0^2 + \frac{2}{3}(t_u - t_l)(\mu - t_0)} \right) \leq \varepsilon. \quad (22)$$

Here, t_0 is the experimentally observed value of T after M measurement rounds, $t_l = \frac{1}{4} - \Delta_\nu - \Lambda_\nu^2$ and $t_u =$

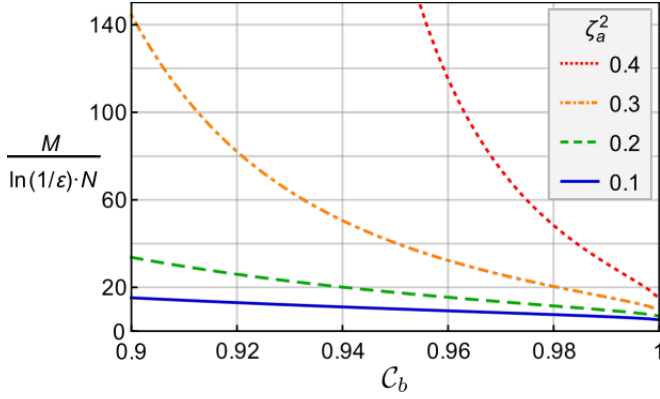


FIG. 3: Number of experimental runs per spins required to rule out non-Bell-correlated states with a confidence of $1 - \epsilon$ as a function of C_b and ζ_a^2 . For $C_b = 0.98$ and $\zeta_a^2 = 0.272$ (as reported in [12]), approximately $17 \cdot \ln(100) \simeq 80$ runs per spin are sufficient to reach a confidence level of 99%.

$\frac{1}{4} + \Delta_\nu + \Lambda_\nu^2(N+1)$ are lower and upper bounds on the random variable \mathcal{T} respectively, and σ_0^2 is its variance for a local state.

We show in Appendix D that the largest p-value is obtained by setting $\mu = 0$ and $\sigma_0^2 = -t_l t_u$. A number of measurements sufficient to exclude the null hypothesis with a probability larger than $1 - \epsilon$ is then given by:

$$M \geq \frac{-2t_l t_u - \frac{2}{3}(t_u - t_l)t_0}{t_0^2} \ln\left(\frac{1}{\epsilon}\right). \quad (23)$$

This quantity can be minimized by choosing the free parameter ν appropriately. As shown in Appendix D, optimizing ν at this stage allows us to reduce the number of measurement by $\sim 30\%$. It is thus clearly advantageous not to consider the witness (18) when evaluating statistical significance.

The number of runs in (23) depends linearly on t_l and

therefore also linearly on N . The ratio $\frac{M}{N}$ thus tends to a constant for large N (see Appendix D for more details). This implies that a number of measurements growing linearly with the system size is both necessary and sufficient to close the statistics loophole [12].

Figure 3 depicts the required number of experimental runs per spin as a function of the scaled collective spin C_b and of the scaled second moment ζ_a^2 . For a confidence level of $1 - \epsilon = 99\%$, between 20 and 500 measurement runs per spins are required in the considered parameter region.

e. Conclusion – In this paper, we introduce a class of multipartite Bell inequalities involving two-body correlators and an arbitrary number of measurement settings. Assuming collective spin measurements, these inequalities give rise to the witness (18), which can be used to determine whether Bell correlations can be detected in a many-body system. This criterion detects states that were not detected by the previously-known witness [12].

We then discuss the statistics loophole arising in experiments involving many-body systems, i.e. the difficulty of ruling out, without further assumptions, non-Bell-correlated states in the presence of finite statistics. We provide a bound, Eq. (23), on the number of measurement rounds that allows one to close this loophole. This bound shows that all non-Bell-correlated states can be convincingly ruled out at the cost of performing a number of measurements that grows linearly with the system size. This opens the way for a demonstration of Bell-correlations in a many-body system free of the statistics loophole.

f. Acknowledgements – We thank Baptiste Allard, Remik Augusiak and Valerio Scarani for helpful discussions. This work was supported by the Swiss National Science Foundation (SNSF) through grants PP00P2-150579, 20020-169591 and NCCR QSIT. NS acknowledges the Army Research Laboratory Center for Distributed Quantum Information via the project SciNet.

Appendix A: Proof of the Bell inequalities

In this appendix, we expand on the proof of Ineq. (2) given in the main text, and cover the case of odd numbers of measurements.

1. Symmetric Bell inequality for m measurements

We consider local measurements on N parties. For each party, one can choose between m measurements $M_k^{(i)}$, where $k \in \{0, 1, \dots, m-1\}$ and $i \in \{1, \dots, N\}$. Each measurement has the two possible outcomes ± 1 . We are interested in Bell inequalities, i.e. inequalities every local theory has to obey [25]. We only consider one- and two-body mean values, so the general form of such an inequality is

$$I_{N,m} = \sum_{k=0}^{m-1} \sum_{i=1}^N \alpha_k^i \langle M_k^{(i)} \rangle + \sum_{k,l} \sum_{i < j} \beta_{kl}^{ij} \langle M_k^{(i)} M_l^{(j)} \rangle \geq -\beta_c, \quad (A1)$$

where $\langle M_k^{(i)} \rangle = \sum_{a \in \{-1,1\}} a \text{Prob}(M_k^{(i)} = a)$ and $\langle M_k^{(i)} M_l^{(j)} \rangle = \sum_{a,b \in \{-1,1\}} ab \text{Prob}(M_k^{(i)} = a, M_l^{(j)} = b)$.

We now restrict ourselves to Bell inequalities which are symmetric under exchange of parties, i.e. $\alpha_k^i = \alpha_k$ and $\beta_{kl}^{ij} = \beta_{kl}$. After defining the symmetrized correlators

$$S_k = \sum_{i=1}^N \langle M_k^{(i)} \rangle, \quad S_{kl} = \sum_{i \neq j} \langle M_k^{(i)} M_l^{(j)} \rangle, \quad (\text{A2})$$

symmetric inequalities can be expressed as

$$I_{N,m} = \sum_{k=0}^{m-1} \alpha_k S_k + \frac{1}{2} \sum_{k,l}^{m-1} \beta_{kl} S_{kl} \geq -\beta_c. \quad (\text{A3})$$

We are interested in cases for which the coefficients are $\alpha_k = m - 2k - 1$ ($k = 0, \dots, m-1$) and $\beta_{k,l} = 1$. We note that $\alpha_{m-k-1} = -\alpha_k$ and claim that local theories have to fulfill the Bell inequalities

$$I_{N,m} = \sum_{k=0}^{m-1} (m - 2k - 1) S_k + \frac{1}{2} \sum_{k,l}^{m-1} S_{kl} \geq - \left\lfloor \frac{m^2 N}{2} \right\rfloor = -\beta_c, \quad (\text{A4})$$

where $\lfloor x \rfloor$ is the largest integer smaller or equal to x .

2. Computation of the local bound

In this section we prove the claim above. One of the most important properties of a local theory is its equivalence to a mixture of deterministic local theory. That is why, by considering only deterministic theories, there is no loss of generality. We can therefore assume that a measurement $M_k^{(i)}$ will lead to an outcome $x_k^i = \pm 1$ with probability 1, i.e. $\langle M_k^{(i)} \rangle = x_k^i$. The two-body correlators S_{kl} can thus be expressed as $S_{kl} = S_k S_l - \sum_{i=1}^N x_k^i x_l^i$. By also taking the antisymmetry of α_k into account, and introducing the quantities

$$A = \sum_{k=0}^{\lfloor \frac{m}{2} \rfloor - 1} (m - 2k - 1) (S_k - S_{m-k-1}), \quad B = \sum_{k=0}^{m-1} S_k, \quad C = \sum_{i=1}^N \left[\sum_{k=0}^{m-1} x_k^i \right]^2. \quad (\text{A5})$$

we arrive at

$$I_{N,m} = A + \frac{1}{2} B^2 - \frac{1}{2} C. \quad (\text{A6})$$

a. Strategy variables

We want to rewrite $I_{N,m}$ further. Therefore we introduce variables counting the strategies chosen by the parties. Because there are m measurements with binary outcomes, the number of possible strategies per party is 2^m . We define the following 2^m variables:

$$\begin{aligned} a_{j_1 < \dots < j_n} &:= \#\{i \in \{1, \dots, N\} | x_k^i = -1 \text{ iff } k \in \{j_1, \dots, j_n\}\} \quad \text{for } n \leq \left\lfloor \frac{m}{2} \right\rfloor, \\ \bar{a}_{j_1 < \dots < j_n} &:= \#\{i \in \{1, \dots, N\} | x_k^i = +1 \text{ iff } k \in \{j_1, \dots, j_n\}\} \quad \text{for } n \leq \left\lfloor \frac{m-1}{2} \right\rfloor, \\ \bar{a}_{j_1, \dots, j_{\lceil \frac{m}{2} \rceil}} &\equiv 0, \end{aligned} \quad (\text{A7})$$

where $\#$ denotes the set cardinality. For example, a_j counts the parties k whose outcomes are $x_k^{j'} = 1 - 2\delta_{jj'}$. \bar{a}_j is the number of parties following the opposite strategy. Variables with n indices thus correspond to a strategy for which either exactly n of the m outcomes are $+1$ or exactly n of the outcomes are -1 , i.e. n outcomes differ from the rest. Note that the conjugate variables in the case of $\frac{m}{2}$ indices are set to zero for the case of even m in order to

prevent strategies from being counted twice. Since every party has to choose one strategy, the variables sum up to N , i.e.

$$a + \bar{a} + \sum_{j=0}^{m-1} (a_j + \bar{a}_j) + \dots = \sum_{n=0}^{\lfloor \frac{m}{2} \rfloor} \sum_{j_1 < \dots < j_n} (a_{j_1 \dots j_n} + \bar{a}_{j_1 \dots j_n}) = N. \quad (\text{A8})$$

Note that S_k can be expressed in terms of the strategy variables as follows:

$$S_k = \sum_{n=0}^{\lfloor \frac{m}{2} \rfloor} \sum_{j_1 < \dots < j_n} (a_{j_1 \dots j_n} - \bar{a}_{j_1 \dots j_n}) y_k^{j_1 \dots j_n}, \text{ with} \quad (\text{A9})$$

$$y_k^{j_1 \dots j_n} = \begin{cases} -1 & \text{if } k \in \{j_1, \dots, j_n\} \\ +1 & \text{else} \end{cases}. \quad (\text{A10})$$

b. Decomposition of A and C in terms of the strategy variables

Equation (A9) results in the following representation of $S_k - S_l$:

$$S_k - S_l = \sum_{n=0}^{\lfloor \frac{m}{2} \rfloor} \sum_{j_1 < \dots < j_n} (a_{j_1 \dots j_n} - \bar{a}_{j_1 \dots j_n}) (y_k^{j_1 \dots j_n} - y_l^{j_1 \dots j_n}). \quad (\text{A11})$$

Clearly, a variable only appears in this expression if $y_k \neq y_l$, i.e. if the strategy is such that the outcome of the k^{th} measurement differs from the l^{th} . This, for example, cannot be the case if the number of indices is zero, i.e. if all measurement outcomes are the same. So a and \bar{a} do not show up in $S_k - S_l$.

With the help of the introduced strategy variables, we can express A as

$$A = \sum_{k=0}^{\lfloor \frac{m}{2} \rfloor - 1} (m - 2k - 1) \sum_{n=0}^{\lfloor \frac{m}{2} \rfloor} \sum_{j_1 < \dots < j_n} (a_{j_1 \dots j_n} - \bar{a}_{j_1 \dots j_n}) (y_k^{j_1 \dots j_n} - y_{m-k-1}^{j_1 \dots j_n}). \quad (\text{A12})$$

and C as

$$\begin{aligned} C &= m^2(a + \bar{a}) + (m-2)^2 \sum_{j=0}^{m-1} (a_j + \bar{a}_j) + (m-4)^2 \sum_{j_1 < j_2} (a_{j_1 j_2} + \bar{a}_{j_1 j_2}) + \dots \\ &= \sum_{n=0}^{\lfloor \frac{m}{2} \rfloor} (m-2n)^2 \sum_{j_1 < \dots < j_n} (a_{j_1 \dots j_n} + \bar{a}_{j_1 \dots j_n}). \end{aligned} \quad (\text{A13})$$

In other words, we notice that if a variable has n indices it contributes to C with a factor $(m-2n)^2$.

c. A bound independent of the number of indices

In this section, we study the contributions of A and C to $I_{N,m}$. For this, we make use of the following theorem:

Theorem 1. *A strategy with n equal outcomes satisfies the inequality*

$$\sum_{k=0}^{\lfloor \frac{m}{2} \rfloor - 1} \left| y_k^{j_1 \dots j_n} - y_{m-k-1}^{j_1 \dots j_n} \right| \leq 2n.$$

Proof. First we note that the summation is such that no $y_k^{j_1 \dots j_n}$ appears twice. Also, we know that since we consider binary outcomes, $|y_k - y_{m-k-1}|$ is either 0 or 2. We thus have

$$\sum_{k=0}^{\lfloor \frac{m}{2} \rfloor - 1} \left| y_k^{j_1 \dots j_n} - y_{m-k-1}^{j_1 \dots j_n} \right| = 2l, \quad l \in \mathbb{N}.$$

Assume now that the above inequality is violated, i.e. $l > n$.

$\Leftrightarrow y_k^{j_1 \dots j_n} \neq y_{m-k-1}^{j_1 \dots j_n}$ for $l > n$ values of k .

\Leftrightarrow The strategy (j_1, \dots, j_n) has l differing outcomes.

This is a contradiction to the definition of the strategy. Therefore the assumption must be wrong and the inequality holds for all strategies. \square

Corollary 1.1. A function $f(k)$ which is monotonically decreasing with k , satisfies

$$\sum_{k=0}^{\lfloor \frac{m}{2} \rfloor - 1} f(k) \left| y_k^{j_1 \dots j_n} - y_{m-k-1}^{j_1 \dots j_n} \right| \leq 2 \sum_{k=0}^{n-1} f(k).$$

Proof. Theorem 1 implies that $y_k \neq y_{m-k-1}$ for at most n values of k . Taking into account that $f(k)$ is monotonically decreasing, we find that

$$\sum_{k=0}^{\lfloor \frac{m}{2} \rfloor - 1} f(k) \left| y_k^{j_1 \dots j_n} - y_{m-k-1}^{j_1 \dots j_n} \right| \leq \sum_{\substack{n \text{ values} \\ \text{of } k}} f(k) \cdot 2 \leq 2 \sum_{k=0}^{n-1} f(k).$$

\square

With the help of Corollary 1.1, we rewrite the quantity A as

$$\begin{aligned} A &= \sum_{n=0}^{\lfloor \frac{m}{2} \rfloor} \sum_{j_1 < \dots < j_n} (a_{j_1 \dots j_n} - \bar{a}_{j_1 \dots j_n}) \sum_{k=0}^{\lfloor \frac{m}{2} \rfloor - 1} \alpha_k \left(y_k^{j_1 \dots j_n} - y_{m-k-1}^{j_1 \dots j_n} \right) \\ &\geq \sum_{n=0}^{\lfloor \frac{m}{2} \rfloor} \sum_{j_1 < \dots < j_n} (a_{j_1 \dots j_n} - \bar{a}_{j_1 \dots j_n}) (-2) \sum_{k=0}^{n-1} \alpha_k = -2 \sum_{n=0}^{\lfloor \frac{m}{2} \rfloor} n(m-n) \sum_{j_1 < \dots < j_n} (a_{j_1 \dots j_n} - \bar{a}_{j_1 \dots j_n}). \end{aligned} \quad (\text{A14})$$

Making use of Eq. (A13) and (7), we then find that

$$\begin{aligned} A - \frac{1}{2}C &\geq \sum_{n=0}^{\lfloor \frac{m}{2} \rfloor} \left[-2n(m-n) - \frac{(m-2n)^2}{2} \right] \sum_{j_1 < \dots < j_n} (a_{j_1 \dots j_n} - \bar{a}_{j_1 \dots j_n}) \\ &= -\frac{m^2}{2} \sum_{n=0}^{\lfloor \frac{m}{2} \rfloor} \sum_{j_1 < \dots < j_n} (a_{j_1 \dots j_n} - \bar{a}_{j_1 \dots j_n}) = -\frac{m^2}{2}N \end{aligned} \quad (\text{A15})$$

d. Putting the pieces together

In order to conclude the proof, we now only miss the contribution of the term B . For this, we look at the case of even and odd m separately. When m is even, $B = \sum_k S_k$ is also even. We thus find that $B^2 \geq 0$. This means that

$$I_{N,m} \geq A - \frac{1}{2}C \geq -\frac{m^2 N}{2}. \quad (\text{A16})$$

If m is odd, B shares the parity of N . That is why we have $B^2 \geq 0$ for even N , and $B^2 \geq 1$ for odd N , resulting in

$$I_{N,m} \geq \begin{cases} -\frac{m^2 N}{2} & \text{for even } N \\ -\frac{m^2 N}{2} + \frac{1}{2} & \text{for odd } N \end{cases}. \quad (\text{A17})$$

In general, the classical bound is thus $\beta_c = \left\lfloor \frac{m^2 N}{2} \right\rfloor$.

Appendix B: Optimization of the witnesses

In this appendix, we optimize the witnesses \mathcal{W}_m as given in Eq. (14) of the main text over the measurement angles. Let us remind the form of \mathcal{W}_m :

$$\mathcal{W}_m = \mathcal{C}_b \sum_{k=0}^{\frac{m}{2}-1} \alpha_k \sin(\vartheta_k) - (1 - \zeta_a^2) \left[\sum_{k=0}^{\frac{m}{2}-1} \cos(\vartheta_k) \right]^2 + \frac{m^2}{4}. \quad (\text{B1})$$

We do this optimization by searching for those angles leading to the minimum of \mathcal{W}_m . This is equivalent to solving the system of equations arising from $\frac{\partial \mathcal{W}_m}{\partial \vartheta_k} = 0$:

$$\frac{\partial \mathcal{W}_m}{\partial \vartheta_k} = (m - 2k - 1) \mathcal{C}_b \cos(\vartheta_k) + 2 \sin(\vartheta_k) (1 - \zeta_a^2) \sum_{l=0}^{\frac{m}{2}-1} \cos(\vartheta_l) = 0. \quad (\text{B2})$$

We eventually want to find angles such that \mathcal{W}_m is negative. To achieve this, the last term of Eq. (B1) must be compensated. Since $\zeta_a^2 \geq 0$, the second term of Eq. (B1) is bounded by $-\frac{m^2}{4}$ and thus cannot be sufficient for a negative \mathcal{W}_m . On the other hand, the first term is bounded by $-\frac{m^2}{4} + 1$ due to the fact that $|\mathcal{C}_b| \leq 1$. So we find that in order to reach $\mathcal{W}_m < 0$, we need $\sin(\vartheta_k)$, $\cos(\vartheta_k)$ and \mathcal{C}_b to differ from zero and $\zeta_a^2 < 1$. In the following studies, we assume these necessary constraints, allowing us to rewrite Eq. (B2) as

$$\frac{2(1 - \zeta_a^2)}{\mathcal{C}_b} \sum_{l=0}^{\frac{m}{2}-1} \cos(\vartheta_l) = -(m - 2k - 1) \frac{\cos(\vartheta_k)}{\sin(\vartheta_k)} \quad \forall k. \quad (\text{B3})$$

Since the left side of Eq. (B3) does not explicitly depend on k , this can only be achieved if both sides are equal to a constant. The assumptions about ζ_a^2 , \mathcal{C}_b and the angles, as reasoned above, allow us to write

$$\frac{2(1 - \zeta_a^2)}{\mathcal{C}_b} \sum_{l=0}^{\frac{m}{2}-1} \cos(\vartheta_l) = \frac{1}{\lambda_m} = -(m - 2k - 1) \frac{\cos(\vartheta_k)}{\sin(\vartheta_k)}, \quad (\text{B4})$$

where λ_m is a constant depending for given \mathcal{C}_b and ζ_a^2 only on m . We find the optimal angles

$$\vartheta_k = -\arctan(\lambda_m(m - 2k - 1)). \quad (\text{B5})$$

For a minimal \mathcal{W}_m , the constants λ_m have to fulfill the self-consistency equations

$$\frac{\mathcal{C}_b}{2\lambda_m(1 - \zeta_a^2)} = \sum_{l=0}^{\frac{m}{2}-1} \cos(\vartheta_l) = \sum_{l=0}^{\frac{m}{2}-1} \frac{1}{\sqrt{1 + \lambda_m^2(m - 2l - 1)^2}}. \quad (\text{B6})$$

Here, we used the fact that $\cos(\arctan(x)) = \frac{1}{\sqrt{1+x^2}}$. For further steps, we note that $\sin(\arctan(x)) = \frac{x}{\sqrt{1+x^2}}$ and define the following functions:

$$\Lambda_m(\lambda_m) := \sum_{k=0}^{\frac{m}{2}-1} \frac{1}{\sqrt{1 + \lambda_m^2(m - 2k - 1)^2}} = \sum_{k=1}^{\frac{m}{2}} \frac{1}{\sqrt{1 + \lambda_m^2(2k - 1)^2}}, \quad (\text{B7})$$

$$\Delta_m(\lambda_m) := \sum_{k=0}^{\frac{m}{2}-1} \frac{\lambda_m(m - 2k - 1)^2}{\sqrt{1 + \lambda_m^2(m - 2k - 1)^2}} = \sum_{k=1}^{\frac{m}{2}} \frac{\lambda_m(2k - 1)^2}{\sqrt{1 + \lambda_m^2(2k - 1)^2}}. \quad (\text{B8})$$

If we assume the representation of the angles given in Eq. (B5), the witnesses can be expressed as

$$\mathcal{W}_m = -\mathcal{C}_b \Delta_m(\lambda_m) - (1 - \zeta_a^2) \Lambda_m^2(\lambda_m) + \frac{m^2}{4} \geq 0, \quad (\text{B9})$$

which holds for non-Bell-correlated states. Note that in this expression, we only assume the *arctan*-angle-distribution, without taking the self-consistency equations into account, i.e. without optimizing the actual differences between angles.

For the case of $m \rightarrow \infty$, we need to rewrite the above witnesses, since \mathcal{W}_m diverges in this limit. We define

$$\mathcal{W}'_m := \frac{\mathcal{W}_m}{m^2} = -\mathcal{C}_b \frac{\Delta_m(\lambda_m)}{m^2} - (1 - \zeta_a^2) \left(\frac{\Lambda_m(\lambda_m)}{m} \right)^2 + \frac{1}{4} \geq 0. \quad (\text{B10})$$

We also have to rewrite the constant λ_m . We define $\lambda_m = \frac{\nu_m}{m}$ and rewrite Eq. (B6) as

$$\frac{\mathcal{C}_b}{2\nu_m(1 - \zeta_a^2)} = \frac{\Lambda_m\left(\frac{\nu_m}{m}\right)}{m}. \quad (\text{B11})$$

If we define $s_k = \frac{2k-1}{m}$, we see that $\frac{1}{m}$ can be expressed as $\frac{s_{k+1} - s_k}{2}$. Note that for $m \rightarrow \infty$, $s_1 \rightarrow 0$ and $s_{m/2} \rightarrow 1$. Using the convention $\nu_\infty = \nu$, we find

$$\begin{aligned} \Lambda_\nu &:= \lim_{m \rightarrow \infty} \frac{\Lambda_m(\nu_m/m)}{m} = \lim_{m \rightarrow \infty} \sum_{k=1}^{\frac{m}{2}} \frac{1}{\sqrt{1 + \nu_m^2 \frac{(2k-1)^2}{m^2}}} \cdot \frac{1}{m} = \lim_{m \rightarrow \infty} \sum_{k=1}^{\frac{m}{2}} \frac{1}{\sqrt{1 + \nu_m^2 s_k^2}} \frac{s_{k+1} - s_k}{2} \\ &= \frac{1}{2} \int_0^1 \frac{1}{\sqrt{1 + \nu^2 s^2}} ds = \frac{\text{arsinh}(\nu)}{2\nu} \end{aligned} \quad (\text{B12})$$

$$\begin{aligned} \Delta_\nu &:= \lim_{m \rightarrow \infty} \frac{\Delta_m(\nu_m/m)}{m^2} = \lim_{m \rightarrow \infty} \sum_{k=1}^{\frac{m}{2}} \frac{\nu_m \frac{(2k-1)^2}{m^2}}{\sqrt{1 + \nu_m^2 \frac{(2k-1)^2}{m^2}}} \cdot \frac{1}{m} = \lim_{m \rightarrow \infty} \sum_{k=1}^{\frac{m}{2}} \frac{\nu_m s_k^2}{\sqrt{1 + \nu_m^2 s_k^2}} \frac{s_{k+1} - s_k}{2} \\ &= \frac{1}{2} \int_0^1 \frac{\nu s^2}{\sqrt{1 + \nu^2 s^2}} ds = \frac{\sqrt{1 + \nu^2}}{4\nu} - \frac{\text{arsinh}(\nu)}{4\nu^2}. \end{aligned} \quad (\text{B13})$$

This yields the witness

$$\mathcal{W}'_\infty = -\mathcal{C}_b \Delta_\nu - (1 - \zeta_a^2) \Lambda_\nu + \frac{1}{4} \quad (\text{B14})$$

$$= -\mathcal{C}_b \left(\frac{\sqrt{1 + \nu^2}}{4\nu} - \frac{\text{arsinh}(\nu)}{4\nu^2} \right) - (1 - \zeta_a^2) \frac{\text{arsinh}^2(\nu)}{4\nu^2} + \frac{1}{4} \geq 0. \quad (\text{B15})$$

We now search for those points in the \mathcal{C}_b - ζ_a^2 -plane that allow for a violation of the correlation witnesses of Ineq. (B9) and (B15). For this purpose, we assume the optimal angles given in Eq. (B6) and (B11) respectively. We define $Z_m(\mathcal{C}_b)$ to be the scaled second moment, as a function of the scaled collective spin, such that \mathcal{W}_m vanishes.

a. $m = 2$

In the case of $m = 2$ measurement settings, we have to solve the following system of equations in order to find Z_2 :

$$0 = \mathcal{W}_2 = -\mathcal{C}_b \frac{\lambda_2}{\sqrt{1 + \lambda_2^2}} - (1 - \zeta_a^2) \frac{1}{1 + \lambda_2^2} + 1, \quad (\text{B16})$$

$$\frac{\mathcal{C}_b}{2\lambda_2(1 - \zeta_a^2)} = \frac{1}{\sqrt{1 + \lambda_2^2}}. \quad (\text{B17})$$

From these, we find the critical line Z_2 and therefore the following condition, satisfied by every non-Bell-correlated state:

$$\zeta_a^2 \geq Z_2(\mathcal{C}_b) = \frac{1}{2} \left(1 - \sqrt{1 - \mathcal{C}_b^2} \right). \quad (\text{B18})$$

b. *Limit $m \rightarrow \infty$*

The critical line Z_∞ is determined by solving the following equation for ζ_a^2

$$0 = \mathcal{W}'_\infty = -\mathcal{C}_b \left(\frac{\sqrt{1 + \nu^2}}{4\nu} - \frac{\text{arsinh}(\nu)}{4\nu^2} \right) - (1 - \zeta_a^2) \frac{\text{arsinh}^2(\nu)}{4\nu^2} + \frac{1}{4}, \quad \text{where } \nu = \sinh \left(\frac{\mathcal{C}_b}{1 - \zeta_a^2} \right). \quad (\text{B19})$$

We find that any non-Bell-correlated state satisfies

$$\zeta_a^2 \geq Z_\infty(\mathcal{C}_b) = 1 - \frac{\mathcal{C}_b}{\text{artanh}(\mathcal{C}_b)}. \quad (\text{B20})$$

Appendix C: Squeezing requirement

Here, we find a bound on the amount of squeezing that is needed as a function of the number of spins N in order to violate the Bell correlation witness (18) described in the main text. Due to the structure of spin systems, the first moment \mathcal{C}_b , the second moment ζ_a^2 and the number of spins N satisfy the following constraints [27]:

$$\zeta_a^2 \geq 1 - \frac{N}{2} \left[\sqrt{(1 - \mathcal{C}_b^2) \left[\left(1 + \frac{2}{N}\right)^2 - \mathcal{C}_b^2 \right]} + \mathcal{C}_b^2 - 1 \right] \quad (\text{C1})$$

(notice that there is an error in the expression given in the reference). For any number of spins N , equating the right-hand side of this constraint with the right-hand side of Eq. (18) gives the maximum value of \mathcal{C}_b under which a violation of the witness (18) is possible. The corresponding maximum value of ζ_a^2 is plotted as a function of the number of spins N in Fig. 2. For large N , this function can be expanded as

$$Z_N^* = 1 - \frac{1}{\omega} - \frac{1}{2\omega^3} - \frac{3}{4\omega^4} - O(\omega^{-5}) \quad (\text{C2})$$

where $\omega = W_{-1}\left(-\frac{1}{2\sqrt{N+1}}\right)$ and W_{-1} is the lower branch of the Lambert W function.

Appendix D: Finite statistics

In this appendix, we introduce four concentration inequalities and determine their bound on the p-value for generic non-Bell-correlated states. We also compare these p-values and choose the optimal one to estimate a number of experimental runs sufficient to exclude non-Bell-correlated states with a confidence $1 - \varepsilon$. Eventually we minimize this number of runs by optimizing the measurement angles.

1. Concentration inequalities

In statistics, concentration inequalities bound the probability that a random variable X exceeds or falls below a certain value. In what follows, we recall the definitions of some of these inequalities. We then discuss some of their properties in view of our problem in the following section.

a. Chernoff bound

The following version of the Chernoff bound was proven by Van Vu at the University of California, San Diego [28]. It was done for discrete, independent random variables. However, the bound also applies in the case of continuous random variables.

Theorem 2. *Let X_1, \dots, X_M be independent random variables with $|X_i| \leq 1$ and expectation values $E[X_i] = 0$ for all*

i. Let $X = \sum_{i=1}^M X_i$ and σ^2 be the variance of X . Then

$$\begin{aligned} P[X \leq -\lambda\sigma] &\leq \exp\left(-\frac{\lambda^2}{4}\right), & \text{for } 0 \leq \lambda \leq 2\sigma, \\ P[X \leq -x_0] &\leq \exp\left(-\frac{x_0^2}{4\sigma^2}\right), & \text{for } 0 \leq x_0 \leq 2\sigma^2. \end{aligned}$$

Corollary 2.1. Let X_1, \dots, X_M be independent random variables with $E[X_i] = \mu$ and $a \leq X_i \leq b$ for all i . Let $X = \frac{1}{M} \sum_{i=1}^M X_i$, $\sigma_i^2 = \text{Var}[X_i]$ and $\sigma_0^2 = \max_i \{\sigma_i^2\}$. Then

$$P[X \leq x_0] \leq \exp\left(-\frac{(\mu - x_0)^2 M}{4\sigma_0^2}\right), \quad \text{for } \mu \geq x_0 \geq \mu - \frac{\sum_i \sigma_i^2}{M(b-a)}.$$

b. Bernstein inequality

The following expression known as Bernstein inequality, was proven by Bernstein in 1927, but we refer to the work of George Bennett [29]. The inequality is valid under certain restrictions for the absolute moments. Since our random variables are bounded, we can be sure these restrictions to be fulfilled.

Theorem 3. Let X_1, \dots, X_M be independent random variables with $E[X_i] = 0$ and $|X_i| \leq \xi$ for all i . Also let $X = \frac{1}{M} \sum_{i=1}^M X_i$ and $\sigma^2 = \frac{1}{M} \sum_{i=1}^M \text{Var}[X_i]$. Then

$$\begin{aligned} P[X \geq x_0] &\leq \exp\left(-\frac{x_0^2 M}{2\sigma^2 + \frac{2}{3}\xi x_0}\right), & \forall x_0 > 0, \\ P[X \leq -x_0] &\leq \exp\left(-\frac{x_0^2 M}{2\sigma^2 + \frac{2}{3}\xi x_0}\right), & \forall x_0 > 0. \end{aligned}$$

Corollary 3.1. Let X_1, \dots, X_M be independent random variables with $E[X_i] = \mu$, $\sigma_i^2 = \text{Var}[X_i]$ and $a \leq X_i \leq b$ for all i . Also let $X = \frac{1}{M} \sum_{i=1}^M X_i$ and $\sigma^2 = \frac{1}{M} \sum_{i=1}^M \sigma_i^2 \leq \sigma_0^2 = \max_i \{\sigma_i^2\}$. Then

$$\begin{aligned} P[X \leq x_0] &\leq \exp\left(-\frac{(\mu - x_0)^2 M}{2\sigma^2 + \frac{2}{3}(b-a)(\mu - x_0)}\right) \\ &\leq \exp\left(-\frac{(\mu - x_0)^2 M}{2\sigma_0^2 + \frac{2}{3}(b-a)(\mu - x_0)}\right), & \forall x_0 < \mu. \end{aligned}$$

c. Uspensky inequality

Uspensky stated in [30] an inequality for a stochastic variable:

Theorem 4. Let X be a random variable with mean value $E[X] = 0$ and $a \leq X \leq b$. Additionally $b \geq |a|$. Let $\sigma^2 = \text{Var}[X]$. Then for $x_0 \leq 0$

$$P[X \leq x_0] \leq \frac{\sigma^2}{\sigma^2 + x_0^2}.$$

Corollary 4.1. Let X_1, \dots, X_M be independent random variables with mean values $E[X_i] = \mu$ and $a \leq X_i \leq b$ for all i . Additionally $b \geq |a|$. Let $\sigma_i^2 = \text{Var}[X_i]$ and $\sigma_0^2 = \max_i \{\sigma_i^2\}$. Then $\sigma^2 = \text{Var}[X] = \text{Var}\left[\frac{1}{M} \sum_{i=1}^M X_i\right] = \frac{1}{M^2} \sum_{i=1}^M \sigma_i^2 \leq \frac{1}{M} \sigma_0^2$ so that for $x_0 \leq \mu$

$$P[X \leq x_0] \leq \frac{\sigma_0^2}{\sigma_0^2 + (x_0 - \mu)^2 M}.$$

d. Berry-Esseen inequality

Andrew C. Berry and Carl-Gustav Esseen proved the following theorem [31]:

Theorem 5. Let X_1, \dots, X_M be independent identically distributed random variables with $E[X_i] = \mu$ for all i and with $X = \frac{1}{M} \sum_{i=1}^M X_i$. Additionally, the variance $\sigma^2 = \text{Var}[X_i] = E[(X_i - \mu)^2]$ and the third absolute moment $\rho = E[|X_i - \mu|^3]$ are finite. Then there exists a constant C such that for all x

$$|F_M(x) - \Phi(x)| \leq \frac{C\rho}{\sigma^3\sqrt{M}}, \quad (\text{D1})$$

where $F_M(x) = P\left[\sqrt{\frac{M}{\sigma^2}}(X - \mu) \leq x\right]$ and $\Phi(x) = \frac{1}{2}(1 + \text{erf}(x))$.

Esseen also proved that the constant C has to fulfill $C \geq \frac{\sqrt{10}+3}{6\sqrt{2\pi}}$. A good estimate for C follows from

$$|F_M(x) - \Phi(x)| \leq \frac{0.33554(\rho + 0.415\sigma^3)}{\sigma^3\sqrt{M}}. \quad (\text{D2})$$

Corollary 5.1. A direct consequence of Theorem 5 is

$$\begin{aligned} P[X \leq x_0] &= P\left[\sqrt{\frac{M}{\sigma^2}}(X - \mu) \leq \sqrt{\frac{M}{\sigma^2}}(x_0 - \mu)\right] \\ &\leq \Phi\left(\sqrt{\frac{M}{\sigma^2}}(x_0 - \mu)\right) + \frac{C\rho}{\sigma^3\sqrt{M}}. \end{aligned}$$

2. Largest p-value of the concentration inequalities

In this section we determine the largest p-value for the concentration inequalities listed above, under the assumption that $x_0 < 0$ and $\mu \geq 0$. We denote with X the random variable $\frac{1}{M} \sum_{i=1}^M X_i$, with $x_l \leq X_i \leq x_u$ and $x_l < 0$.

The crucial point here is that we have no information about the random variable apart from its non-negative mean value. This lack of information directly disqualifies the Berry-Esseen inequality of Corollary 5.1 as a potential tight bound. Indeed, the Berry-Esseen bound does not result in a tighter restriction than the trivial bound $P[X \leq x_0] \leq 1$. To see this, we consider the following distribution with three peaks:

Consider the case, for which all X_i satisfy the following probability distribution:

$$P[X_i = x] = \begin{cases} p_l & \text{for } x = x_l \\ p_u & \text{for } x = x_u \\ 1 - p_l - p_u & \text{for } x = 0 \\ 0 & \text{else} \end{cases}, \quad (\text{D3})$$

where $p_l, p_u < 1$. Additionally, we demand $E[X_i] = 0$ leading to the condition $-p_l x_l = p_u x_u$. We thus arrive at the variance and third absolute moment:

$$\sigma^2 = p_u x_u (x_u - x_l), \quad (\text{D4})$$

$$\rho = p_u x_u (x_u^2 + x_l^2) \quad (\text{D5})$$

$$\Rightarrow \frac{\rho}{\sigma^3} = \frac{x_u^2 + x_l^2}{\sqrt{p_u x_u (x_u - x_l)^3}}. \quad (\text{D6})$$

In this expression, p_u remains as a parameter scaling the weight on the edges compared to the weight at $x = 0$. Note that for $p_u \rightarrow \frac{-x_l}{x_u - x_l}$, we arrive at a binomial distribution while for $p_u \rightarrow 0$ we have a delta distribution. From Eq. (D6), we see that in the limit $p_u \rightarrow 0$, $\frac{\rho}{\sigma^3} \rightarrow \infty$. Thus, we can write for the Berry-Esseen bound

$$P[X \leq x_0] \leq \Phi\left(\sqrt{\frac{M}{\sigma^2}}(x_0 - \mu)\right) + \frac{C\rho}{\sigma^3\sqrt{M}} \leq \frac{C\rho}{\sigma^3\sqrt{M}} = \frac{C}{\sqrt{M}} \frac{x_u^2 + x_l^2}{\sqrt{p_u x_u (x_u - x_l)^3}} \xrightarrow[p_u \rightarrow 0]{} \infty. \quad (\text{D7})$$

Since we have no information on the actual probability distribution, the largest p-value of the Berry-Esseen bound is 1. We thus restrict our interests to the Chernoff, Bernstein and Uspensky bounds.

The inequalities of Corollaries 2.1, 3.1 and 4.1 have certain properties in common: They all depend on the variance σ_0^2 and the mean value μ . More explicitly, the dependence on σ_0^2 in all three cases is such that if one increases the variance, the bounds are also increased. We therefore use the following strategy to determine the largest p-values: We increase the variance of an arbitrary random variable in a way leaving the mean value unaffected. We eventually arrive at an easy-to-handle probability distribution with a maximal variance. The bounds resulting from this distribution then serve as upper bounds for all distributions of the same mean value. The bounds given by the concentration inequalities will then only depend on the mean value, so that we can optimize over μ .

Eventually, we show that the largest p-value results from the binomial distribution centered around $\mu = 0$.

Theorem 6. *Let X be a random variable in the interval $[a, b]$ with an arbitrary probability distribution and with $E[X] = \mu$. Let X_{bi} be a binomially distributed random variable with peaks at the edges a and b . Furthermore, X_{bi} has the same mean value as X , i.e. $P[X_{bi} = a] = \frac{\mu-b}{a-b}$, $P[X_{bi} = b] = \frac{a-\mu}{a-b}$ and $P[X_{bi} = x] = 0$ otherwise. Additionally, let $a \leq 0$. Then*

$$\text{Var}[X_{bi}] \geq \text{Var}[X].$$

Proof. We define a third random variable Y satisfying

$$P[Y = x] = \begin{cases} P[X = x] & \text{if } x \notin dx_0 \cup \{a, b\} \\ 0 & \text{if } x \in dx_0 \\ P[X = a] + qP[X \in dx_0] & \text{if } x = a \\ P[X = b] + (1-q)P[X \in dx_0] & \text{if } x = b \end{cases}. \quad (\text{D8})$$

Here, dx_0 is an infinitesimal set around $x_0 \in]a, b[$. So in other words, the probability distribution function of Y is almost the same as the one of X . The only difference is that the set dx_0 is "cut out" and the probabilities at a and b are increased (see Fig. 4). They are increased in a way which leaves the mean value unaffected, i.e. q is chosen such that $E[Y] = E[X] = \mu$:

$$\begin{aligned} E[Y] &= E[X] - P[X \in dx_0]x_0 + qP[X \in dx_0]a + (1-q)P[X \in dx_0]b \\ &= \mu + P[X \in dx_0](-x_0 + q(a-b) + b) = \mu \\ \Leftrightarrow \quad q &= \frac{x_0 - b}{a - b}. \end{aligned} \quad (\text{D9})$$

With this and $a \leq 0$, we show that $\text{Var}[Y] \geq \text{Var}[X]$:

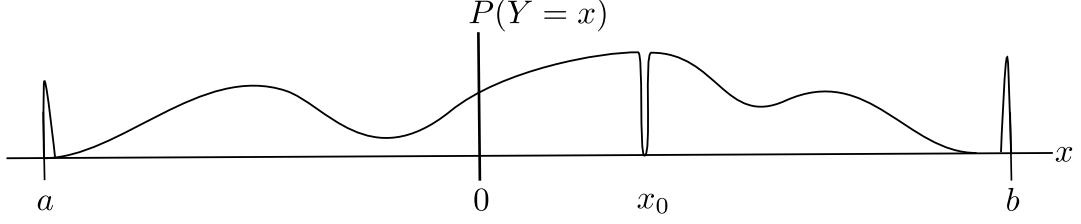
$$\begin{aligned} \text{Var}[Y] &= \text{Var}[X] + P[X \in dx_0] [-x_0^2 + q(a^2 - b^2) + b^2] = \text{Var}[X] + P[X \in dx_0] [-x_0^2 + x_0(a+b) - ab] \\ &\geq \text{Var}[X] + P[X \in dx_0] [-x_0^2 + x_0(a+x_0) - ax_0] \geq \text{Var}[X] + P[X \in dx_0] [-x_0^2 + x_0^2] = \text{Var}[X]. \end{aligned} \quad (\text{D10})$$

So we find that $\text{Var}[Y] \geq \text{Var}[X]$, with the equal sign only if $P[X \in dx_0] = 0$. By induction, one can gradually "cut out" all the other points in $]a, b[$. During this process, the variance is constantly increased while the mean value remains unchanged. Eventually, one arrives at the random variable X_{bi} with its binomial distribution. \square

By applying the different bounds and using Theorem 6, we find

$$P[X \leq x_0] \leq \begin{cases} \exp\left(-\frac{(\mu-x_0)^2 M}{4\sigma_{bi}^2(\mu)}\right) =: p_C(\mu, M) & \text{Chernoff} \\ \exp\left(-\frac{(\mu-x_0)^2 M}{2\sigma_{bi}^2(\mu) + \frac{2}{3}(x_u-x_l)(\mu-x_0)}\right) =: p_B(\mu, M) & \text{Bernstein} \\ \frac{\sigma_{bi}^2(\mu)}{\sigma_{bi}^2(\mu) + (x_0-\mu)^2 M} =: p_U(\mu, M) & \text{Uspensky} \end{cases}, \quad (\text{D11})$$

where $\sigma_{bi}^2(\mu) = (x_u - \mu)(\mu - x_l)$. As stated above, $\mu \in [0, x_u]$. Restricted to this interval, the three functions p_C , p_B and p_U are strictly monotonous decreasing with μ and therefore take their maximal values at $\mu = 0$. This allows us

FIG. 4: Sketch of the probability distribution function of Y in Theorem 6.

to write

$$P[X \leq x_0] \leq \begin{cases} \exp\left(-\frac{x_0^2 M}{4\sigma_{bi}^2}\right) =: p_C(M) & \text{Chernoff} \\ \exp\left(-\frac{x_0^2 M}{2\sigma_{bi}^2 - \frac{2}{3}(x_u - x_l)x_0}\right) =: p_B(M) & \text{Bernstein} \\ \frac{\sigma_{bi}^2}{\sigma_{bi}^2 + x_0^2 M} =: p_U(M) & \text{Uspensky} \end{cases} \quad (\text{D12})$$

with $\sigma_{bi}^2 := \sigma_{bi}^2(0) = -x_l x_u$.

We identify M with the number of experimental runs, and determine the minimum number of runs required to have $P[X \leq x_0] \leq \varepsilon$. This corresponds to solving $p_i(M) \leq \varepsilon$ to M , where $i = C, B, U$. We find

$$M \geq \begin{cases} M_C := \frac{4\sigma_{bi}^2}{x_0^2} \ln\left(\frac{1}{\varepsilon}\right) & \text{Chernoff} \\ M_B := \frac{2\sigma_{bi}^2 - \frac{2}{3}(x_u - x_l)x_0}{x_0^2} \ln\left(\frac{1}{\varepsilon}\right) & \text{Bernstein} \\ M_U := \frac{\sigma_{bi}^2(1-\varepsilon)}{\varepsilon x_0^2} & \text{Uspensky} \end{cases} \quad (\text{D13})$$

3. Comparison of the bounds

We now compare the three bounds stated in Ineq. (D13) and show that in the regime of interest, the Bernstein bound is the tightest bound. By comparing M_B to M_C , one finds

$$\begin{aligned} M_B \cdot \frac{x_0^2}{\ln(1/\varepsilon)} &= 2\sigma_{bi}^2 - \frac{2}{3}(x_u - x_l)x_0 < 2\sigma_{bi}^2 - (x_u - x_l)x_0 \\ &< -2x_l x_u - x_u x_l - x_l x_u = -4x_l x_u = 4\sigma_{bi}^2 = M_C \cdot \frac{x_0^2}{\ln(1/\varepsilon)} \\ &\Rightarrow M_B < M_C. \end{aligned} \quad (\text{D14})$$

This means the bound resulting from Bernstein's inequality is tighter than the Chernoff bound for all values of x_l , x_u and x_0 .

Now we compare p_B to p_U . Plotting both functions reveals that Uspensky's inequality is better for small numbers of experimental runs, whereas Bernstein's is better for larger numbers of runs. We now want to estimate for which ε both approximations require the same number of runs M . We therefore set $M_B = M_U$:

$$\frac{2\sigma_{bi}^2 - \frac{2}{3}(x_u - x_l)x_0}{x_0^2} \ln\left(\frac{1}{\varepsilon}\right) = \frac{\sigma_{bi}^2(1-\varepsilon)}{\varepsilon x_0^2} \Leftrightarrow 2 + \frac{2}{3} \frac{(x_u - x_l)(-x_0)}{\sigma_{bi}^2} = \frac{1-\varepsilon}{\varepsilon \ln\left(\frac{1}{\varepsilon}\right)}. \quad (\text{D15})$$

The function on the right side of Eq. (D15) is strictly monotonous decreasing with ε . The left side is just a number. So the minimal ε possible is achieved if the left side is maximized. We thus make the following estimation:

$$\frac{(x_u - x_l)|x_0|}{\sigma_{bi}^2} = \frac{(x_u + |x_l|)|x_0|}{|x_l|x_u} \leq \frac{x_u|x_l| + |x_l|x_u}{|x_l|x_u} = 2.$$

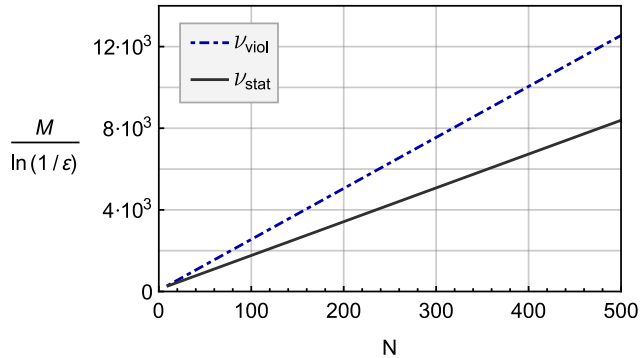


FIG. 5: Plot of the required number of runs as a function of the number of atoms, with $C_b = 0.98$ and ζ_a^2 . The blue dash-dotted line results from the ν which maximizes the violation of witness (19) whereas the solid black line corresponds to the ν minimizing the number of runs. As one can see, the optimization of ν in terms of statistics reduces the number of runs by a factor of approximately $\frac{2}{3}$.

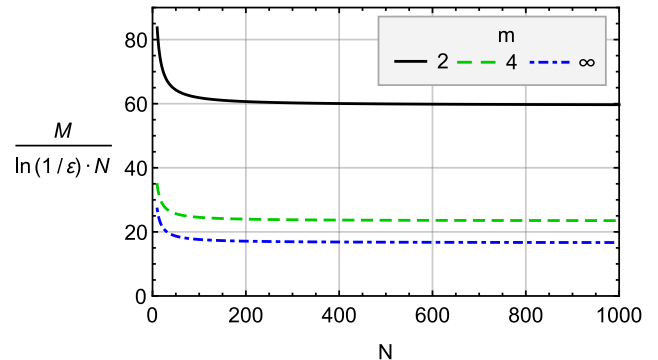


FIG. 6: Plots of the required number of experimental runs M per spin for a confidence of ε in units of $\ln(\frac{1}{\varepsilon})$, as a function of the number of spins N . The plots are for $m = 2, 4, \infty$, with $\zeta_a^2 = 0.272$ and $C_b = 0.98$. The ratios tend to constants for larger systems.

For this value, Eq. (D15) yields $\varepsilon \approx 0.127$. This means that the Uspensky bound can only be better than the Bernstein bound for $\varepsilon \geq 0.127$. Since we are interested in probabilities $\varepsilon \leq 0.1$, the Bernstein bound remains the preferred one.

4. Statistical optimization

Let us now present the result of numerical studies on the number of measurement runs allowing one to rule out non-Bell-correlated states. Here we rely on the Bernstein bound presented in Ineq. (D13).

a. Choice of settings

First, we study the choice of measurement settings, i.e. the choice of ν in Eq. (B15), which allow for the strongest statistical claim. For this, we minimize the number of experimental runs sufficient to conclude about the presence of a Bell-correlated state with given confidence $1 - \varepsilon$ over the choices of ν . We then compare this number of measurements M to the one obtained when choosing the $\nu = \sinh\left(\frac{C_b}{1-\zeta_a^2}\right)$, i.e. for the settings which maximize the witness value (see Eq. (B19)).

The result of this comparison is presented in Figure 5 for a particular choice of C_b and ζ_a^2 . Clearly, it is advantageous to reoptimize the measurement setting in order to maximize the statistical evidence, and thus a different witness should be considered in this case. In the rest of this appendix, measurement settings are optimized in order to maximize statistical evidence.

b. Linear relation between the number of measurement runs and the number of spins

Figure 6 depicts the number of measurement runs per spin needed in order to reach a confidence of $1 - \varepsilon$. The plots are for the values of C_b and ζ_a^2 stated in [12], and for $m = 2, 4, \infty$ settings. As explained in the main text, we observe that the ratio $\frac{M}{N}$ tends to a constant as N increases. This plot also illustrates the gain obtained in using the newly derived Bell inequality with m settings.

c. Minimum squeezing requirement for finite M

As discussed in the main text, violation of our witness requires a finite amount of squeezing for systems of finite size N . Here we study this requirement when finite number of measurement runs are taken into account. In Fig. 7 we show the upper bound on ζ_a^2 as a function of the number of spins N and the number of measurement runs M for

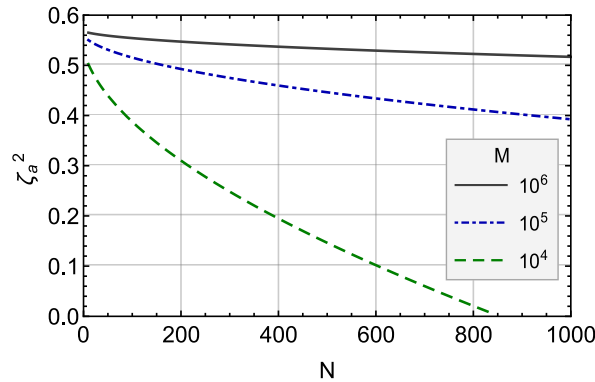


FIG. 7: Plots of the required scaled second moment ζ_a^2 as a function of the number of atoms N , with $C_b = 0.98$ and $\varepsilon = 0.1$. The different curves represent different numbers of experimental runs ($M = 10^4, 10^5, 10^6$).

the case $C_b = 0.98$, $\varepsilon = 0.1$. Although Fig. 2 in the main text shows that a violation with $N = 1000$ spins is possible when $\zeta_a^2 \lesssim 0.8$, we see here that a squeezing of $\zeta_a^2 \lesssim 0.5$ is needed if the number of measurement is limited to 10^6 .

-
- [1] V. Scarani, Acta Physica Slovaca 62, 347 (2012).
 - [2] D. C. Bailey, Royal Society Open Science 4, 160600 (2017).
 - [3] W. J. Youden, 14, 111 (1972).
 - [4] J-D. Bancal, N. Gisin, Y-C. Liang, S. Pironio, Phys. Rev. Lett. 106, 250404 (2011).
 - [5] J. T. Barreiro, J-D. Bancal, P. Schindler, D. Nigg, M. Hennrich, T. Monz, N. Gisin, R. Blatt, Nature Physics 9, 559 (2013).
 - [6] B.P. Lanyon, M. Zwerger, P. Jurcevic, C. Hempel, W. Dür, H.J. Briegel, R. Blatt, and C.F. Roos, Phys. Rev. Lett. 112, 100403 (2014).
 - [7] M. Eibl, S. Gaertner, M. Bourennane, C. Kurtsiefer, M. Żukowski, H. Weinfurter, Phys. Rev. Lett. 90, 200403 (2003).
 - [8] Z. Zhao, T. Yang, Y-A. Chen, A-N. Zhang, M. Żukowski, J-W. Pan, Phys. Rev. Lett. 91, 180401 (2003).
 - [9] J. Lavoie, R. Kaltenbaek, K. J. Resch, New J. Phys. 11, 073051 (2009).
 - [10] B. Hensen, H. Bernien, A.E. Drau, A. Reiserer, N. Kalb, M.S. Blok, J. Ruitenber, R.F.L. Vermeulen, R.N. Schouten, C. Abelln, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D.J. Twitchen, D. Elkouss, S. Wehner, T.H. Taminiau, R. Hanson, Nature 526, 682 (2015); L. K. Shalm et al., Phys. Rev. Lett. 115, 250402 (2015); M. Giustina et al., Phys. Rev. Lett. 115, 250401 (2015); W. Rosenfeld, D. Burchardt, R. Garthoff, K. Redeker, N. Ortégel, M. Rau, H. Weinfurter, arXiv:1611.04604
 - [11] J. Tura, R. Augusiak, A. B. Sainz, T. Vértesi, M. Lewenstein, A. Acín, Science, 344, 1256 (2014).
 - [12] R. Schmied, J-D. Bancal, B. Allard, M. Fadel, V. Scarani, P. Treutlein, N. Sangouard, Science, 352, 6284 (2016).
 - [13] S. Pelisson, L. Pezzè, A. Smerzi, Phys. Rev. A 93, 022115 (2016).
 - [14] R. Gill, arXiv:1207.5103
 - [15] Y. Zhang, S. Glancy, E. Knill, Phys. Rev. A 84, 062118 (2011).
 - [16] C. Gross, T. Zibold, E. Nicklas, J. Estève, M. K. Oberthaler, Nature 464, 1165 (2010).
 - [17] M. F. Riedel, P. Böhi, Y. Li, T. W. Hänsch, A. Sinatra, P. Treutlein, Nature 464, 1170-1173 (22 April 2010).
 - [18] Bernd Lücke, J. Peise, G. Vitagliano, J. Arlt, L. Santos, G. Tóth, C. Klempt, Phys. Rev. Lett. 112, 155304 (2014).
 - [19] L. Pezzè, A. Smerzi, M. K. Oberthaler, R. Schmied, P. Treutlein, Non-classical states of atomic ensembles: fundamentals and applications in quantum metrology (2016), arXiv preprint arXiv:1609.01609.
 - [20] A. S. Sørensen, L-M. Duan, J. I. Cirac, P. Zoller, Nature 409, 63 (2001).
 - [21] P. Hyllus, L. Pezzè, A. Smerzi, G. Tóth, Phys. Rev. A 86, 012337 (2012).
 - [22] G. Tóth, C. Knapp, O. Gühne, H. J. Briegel, Phys. Rev. A 79, 042334 (2009).
 - [23] J. Tura, A. B. Sainz, T. Vértesi, A. Acín, M. Lewenstein, R. Augusiak, Journal of Physics A 42, 424024 (2014).
 - [24] J. Tura, G. de las Cuevas, R. Augusiak, M. Lewenstein, A. Acín, J. I. Cirac, arXiv:1607.06090.
 - [25] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, S. Wehner, Rev. Mod. Phys. 86, 419 (2014).
 - [26] J-D. Bancal, N. Gisin, S. Pironio, J. Phys. A 43, 385303 (2010).
 - [27] A. S. Sørensen, K. Mølmer, Phys. Rev. Lett. 86, 4431 (2001).
 - [28] <http://cseweb.ucsd.edu/~klevchen/techniques/chernoff.pdf>
 - [29] G. Bennett, Journal of the American Statistical Association, 57, 297 (1962).
 - [30] J. V. Uspensky, Introduction to Mathematical Probability. No. 519.1 U86. 1937.
 - [31] A. C. Berry, Trans. Amer. Math. Soc., 49, 122-136 (1941).